**HIPPI-6400** X3T11 ad hoc HIPPI Special Working Group / HNF Technical Committee **HIPPI-6400**
Working Meeting to Address HIPPI-6400 Issues Only
Tuesday through Thursday, March 4-6, 1997
San Jose, CA

## 1. Opening remarks and introductions

The Chairman, Don Tolmie of Los Alamos National Laboratory, opened this HIPPI-6400 meeting and thanked Barbara Weber and Berg Electronics for hosting this meeting. This group is constituted as both the HIPPI special working group (SWG) under X3T11, and the HIPPI Networking Forum (HNF) - Technical Committee (TC).

Don then lead a round of introductions. The list of attendees is at the end of these minutes.

No one objected to including their e-mail address and phone number in the list of attendees that are part of the minutes. It was requested that Don Tolmie provide an updated HIPPI attendees list for distribution.

## 2. Review / modify the draft agenda

The draft agendas were available on the web prior to the meeting. Hard copies were available at the meeting. It was agreed to swap the order of items, putting HIPPI-ST early in the agenda to take advantage of some SGI experts. HIPPI-6400-SC, and its ARP and broadcast issues, were set to start at 8 AM on Wednesday to make sure that we got through them. Don Tolmie volunteered to take the meeting minutes. These minutes are in the agenda order, which did not exactly match the sequence that items were covered during the meeting.

## 3. Review minutes of previous meetings

### 3.1 February 3-4, 1997, San Jose, CA

The February minutes were reviewed; no corrections or changes were suggested. Joe Parker moved, and Roger Ronald seconded, to accept the February minutes as written. Passed unanimously.

### 3.2 Review action items from February meeting

1. Fred Templin, Jeff Young, and Greg Chesson to begin an IP / ARP over HIPPI-6400 RFC. (Carryover)
2. Greg Chesson to register the EtherTypes with INA / Xerox. (Done)
3. Greg Chesson to have Art Beckman look into getting a 24-bit organization identifier for HNF's use in ULAs. (Carryover)
4. Greg Chesson to present text to reflector to describe RTR setup using the Persistent bit. (Carryover)
5. Henry Brandt and Hansel Collins to collect values for completion of copper interface specifications. (Done)
6. Greg Chesson and Don Sanders to have SGI review the connector layout as it relates to the SuMAC chip. (Carryover)
7. Greg Chesson and Don Sanders to check all error states, e.g., what happens if a HIPPI-ST command is corrupted or lost. (Done)
8. Craig Davidson to add text to 4.10 of PH describing the 16-bit optical interface. (Overcome by events)
9. Don Tolmie to correct the January meeting minutes on the HIPPI web page. (Done)
10. John Ellis to send the part number for the 100-pin Berg connector and receptacle to Don Tolmie for inclusion in the document. (Done)
11. Fred Templin to provide accurate reference in 7.2 of PH for "assigned numbers". (Done)
12. Michael McGowen to review 8.2.1 of SC and present a fix when two directly connected end points try to use the same ULA. (Carryover)
13. Greg Chesson to draft initial text for Block acknowledge / retransmission functions. (Carryover)
14. Greg Chesson to compare putting all Operation rejects in State_Response's vs. in expected responses, e.g., Request_Port_Response with Reject = 1. (Done)
15. Greg Chesson to investigate how long an Originating Source should keep buffers for possible retransmission. (Combined with action item 13)
16. Don Tolmie to draft a Project Proposal and initial HIPPI-6400-OPT draft document. (Carryover)

17. James Hoffman to finish revising ST Annex C for the next revision of ST.
18. Roger Ronald to update HIPPI-6400-SC Rev 0.8 with changes agreed to at the February meeting. (Done)
19. Don Tolmie to update HIPPI-6400-ST Rev 0.3 with the changes agreed to at the February meeting. (Done)
20. Don Tolmie to update HIPPI-6400-PH Rev 1.0 with the changes agreed to at the February meeting. (Done)

## 4. HIPPI-6400-PH

### 4.1 Review changes from Rev 0.1 → Rev 1.1

The Abstract, Introduction, Foreword, and Scope all changed to reflect moving the optical interface to another document. This change also rippled through other places in the document. Most of the text was accepted as changed, but some re-wording of 4.10 was requested. Some of the "optical" comments in 13.1 will also be removed.

The definition of ULA was harmonized with HIPPI-6400-SC. The question of whether or not "MAC" should be included in "Universal LAN MAC Address" was resolved by Fred Templin doing some web searches for both ways (and both turned up). It was decided to include "MAC" in the definition, but omit it in the acronym.

A rewording for VC3 size in 6.2 was suggested, i.e., make it more consistent with the other VC size wordings.

During the review of IP over HIPPI-6400-PH (see 4.2 below) it was decided that a high-level user should not be required to include the pad at the end of a message, i.e., pad if the message did not fill an integral number of 32-byte micropackets. Hence, the requirement for supplying the pad was moved to HIPPI-6400-PH, and to avoid creating a covert channel the pad was required to be zeros. SGI had resisted this when it was first brought up many months ago, but now, with the addition of their chip doing HIPPI-ST, have the capability to include the pad.

The training sequences had been changed so that the training sequence for an 8-bit system is the same as for a 16-bit system. An error was noted in the CLOCK waveform for 8-bit system training sequences and will be corrected. The 100 ms "Training" timer was renamed to the "Deadman" timer.

A considerable number of changes were made to the reset and initialize text at the request of Dave Parry and others from SGI. During the review in the meeting, even more changes were suggested, including changes to figure 16.

Other editorial changes were reviewed, some suggested by Dean Liberty of IBM, Tom Gilbert of Harris, and others, by e-mail. The document is getting better, and the comments are appreciated. For example, they pointed out that nowhere in the document had it said what the limit was on the number of unacknowledged micropackets; this has now been remedied by the addition of wording in 6.4.

### 4.2 IP over HIPPI-6400-PH

Craig Davidson presented a proposed normative annex for HIPPI-6400-PH describing how to carry IP datagrams over HIPPI-6400-PH. The comment about carrying all IP traffic on VC1 got the juices going, and it was resolved that VC1 would be the default, but other VCs could be used, albeit carefully. Also, as previously noted, requiring pad to be inserted at this level was dropped and moved to HIPPI-6400-PH. It was also decided that this proposed normative annex belongs in the proposed RFC for IP over HIPPI-6400, not as an annex to HIPPI-6400-PH (Fred Templin said that this proposal would be included in the RFC that he is actioned to do). ARP for HIPPI-6400-PH should also be in an RFC.

### 4.3 HIPPI-6400 to HIPPI-800 IP Translation

Craig Davidson presented a proposed normative annex for HIPPI-6400-PH describing the translation between mapping IP between HIPPI-6400-PH and HIPPI-800. Michael McGowen noted that there are differences based on whether the hosts are attached via a router function, or are "channel attached". It may require a proxy ARP server on the HIPPI-800 side. The only good reason for the LE header is to support ARP. It might be better to ignore existing IP over HIPPI-800 formats and create a new format (omitting the LE header) which looks like the HIPPI-6400 header with an I-Field and FP header attached.

The M-Type field in the LE header will be filled with zeros, and more text will be added to describe the router/channel-attached environments.

### 4.4 Copper issues

Hansel Collins had provided new electrical specifications, which Don Tolmie integrated into HIPPI-6400-PH Rev 1.1. The new specifications were reviewed. Hansel, and Henry Brandt of IBM, were thanked for their work in providing the format and raw numbers for the electrical specifications. The format changed from text descriptions to tables, and many parameters were added in this revision. The detailed review of clauses 15 (timing) and 16 (copper cable interface) pointed out errors, things that should be removed, and things that should be added. For example, the 8-bit system column for the copper cable, and the jitter, will be removed from table 8.

A local electrical interface to drive optical transceivers was also added in this revision, and it in turn needs additional changes. A timing table, similar to table 8 will need to be added, i.e., it is the 8-bit column from the original table 8.

With the equalizer optimized for 50 meters, and on the board, the question of how short a cable could be was raised, but not answered. People want to avoid large coils of cable under the floors as the alternative to short cables. The concept of moving the equalizers to the backshell was re-opened; we had considered it before but then were looking more at active (rather than passive) equalizers and they took up too much room. John Ellis and Herb Van Deusen took an action item to investigate the possibility of putting the passive equalizers in the backshell, with a goal of making a decision by April.

The connector drawings were reviewed, with the major error being the dimensions for B6 and B8 were reversed. The A16 dimension in the table will be replaced with a note that the dimension is dependent on board thickness, i.e., what was there originally was for 0.0625" board.

### 4.5 Service Interface

Don Tolmie had hoped to have a draft of the HIPPI-6400-PH service interface for review at this meeting, but was unable to complete the draft.

## 5. HIPPI-6400-SC

### 5.1 Review changes from Rev 0.8 → Rev 0.9

Roger Ronald lead a discussion of the changes in HIPPI-6400-SC. HIPPI-6400-SC was changed this month to provide broadcast support. This was a contentious topic and there were some strong opinions expressed during the review. Discussions also touched on what should be in the HIPPI-6400-SC document and what should be in RFCs

The major area of concern was that use of non-standard addresses for standard functions would create problems. It was proposed that the ability to modify the port output mapping table for the broadcast server would resolve this issue. The E-Systems switch under development would need to support rapid port table changes to implement this; second-generation switches would probably support broadcast as a native capability. It was suggested that the text be broken into "Broadcast Capabilities" and "Broadcast Emulation" sections.

### 5.2 ARP / broadcast scheme

Fred Templin presented some continuing work on ARP and broadcast for HIPPI-6400 switches which generated a lot of discussion and a lot of heat. The goals that were stated by the group were:

- Use standard methods and procedures where possible, e.g., 802.1d.
- Consider an accelerated broadcast function (server) for a non-broadcast-capable switch.
- Design the accelerated broadcast so that it has minimum impact on second-generation switches.
- Minimize perturbations to the switch hardware currently under development.

Fred provided a picture of a configuration that was used as the basis for the discussions, and agreed to make an electronic copy available so that the discussions can continue over e-mail.

It was seen that broadcast is a key attribute needed to make ARP work across heterogeneous fabrics.

## 6. HIPPI-ST

### 6.1 Review changes from Rev 0.3 → Rev 0.4

The changes to make the new revision were reviewed. Many minor edits were accepted. It was agreed to add definitions for "Data Operation" and "log".

The size of Data Operations was changed so that the data payload is now $2^{31}$ bytes, i.e., get rid of all of the funny numbers like –64, or –9, etc. We could closely specify a number for when HIPPI-6400-PH was the lower layer, but could not be sure what the fudge factor would be for all other media. Hence, $2^{31}$ is safe for all cases, and is still a very large number. This affects the text in 4.2, 6.2, 8.5, and figures 3 and A.1.

It was suggested to change figure 4 with "Max-STU size" replacing "STU-Size", and to use multiple non-horizontal arrows between the "Block Descriptor" and "Buffer Descriptor Table". The text describing the buffer sizes as integral powers of 2 will be made uniform throughout the document. Text will be added to 4.4.9 about using the maximum values for a parameter to achieve best performance.

A major change was to revert back to rejecting operations with the Reject bit in the normal response rather than using a State_Response. A table was included in 4.5.3 to list the responses to rejected operations, and the text for the individual operations was updated. The change was made at the request of the chip implementers – they felt that it was easier to decode this way. With this change in place, and considering all of the other changes made recently, Jeffrey Chung took an action item to verify that the summary tables 2 and 3 accurately reflect the text for each of the Operations. At the e-mail request of Dean Liberty of IBM, some explanatory text was added at the front of each Operation in clause 8, giving a simple explanation of what the Operation was about.

Greg Chesson requested a change to the Notify flag bit, changing its polarity so that it was more of a Mute or Suppress for a Data Operations. Greg noted that the Notify action was the default on Control Operations, and it was now somewhat confusing as to whether Notify should be in all Control Operations or not – the change should clarify it. Greg took an action item to propose new text.

Quite a bit of time was spent going over the error processing text in clause 9, and many changes were agreed to for the next revision. It was noted that lots of errors cause a Reject, and the possibility of adding a "reason code" to help the Reject recipient figure out why it was rejects was discussed. Greg Chesson and Jeffrey Chung took an action item to further consider a "reason code".

### 6.2 HIPPI-ST over HIPPI-FP (proposed annex)

Craig Davidson presented a proposed normative annex for HIPPI-ST describing how to carry ST over HIPPI-FP. One issue that was raised previously was the use of "short burst last". Cray has IOS code for IPI-3 that uses short burst first – special support for this was not deemed necessary. With HIPPI-6400, Cray will use -ST, and short burst last. No other short first burst cases have been identified.

Craig's proposal was modified by the deletion of the HIPPI-FP D1_Area; now everything will be placed in the D2_Area. The D2_Size parameter will equal M_len + 16. We also decided that the most-significant 24 bits of the addresses will be the HNF Org number (Chesson is acquiring), the next 12 bits will be 1's (x'FFF'), and the low-order 12 bits will be the HIPPI-800 address.

### 6.3 Translating Scheduled Transfer Protocol between HIPPI-6400 and HIPPI-FP (proposed annex)

Craig Davidson presented a proposed normative annex for HIPPI-ST describing how to translate the ST protocol between HIPPI-6400 and HIPPI-FP. It was agreed that the portion of the proposal addressing fragmenting between systems was not appropriate for the standard as it seemed to apply mostly to existing Cray machines that do not have the ability to assemble / disassemble large blocks without incurring a high interrupt overhead. Craig took an action item to update the proposal with the changes agreed to at this meeting.

## 7. HIPPI-6400 MIB

Von Welch of NCSA had not updated his HIPPI-6400 MIB since HIPPI-6400-PH Rev 0.9. Some people have expressed some concern over the items listed, and Von took an action item to contact as many of the HIPPI-6400 MIB developers and users

as possible to discuss their concerns. Von also took an action item to develop a presentation on the MIB for a future meeting.

## 8. Other items

None

## 9. Future meeting schedule

### 9.1 April 7-8, Palm Springs, CA

During the X3T11 April plenary week , the following HIPPI meetings are scheduled:

Monday, April 7 -
1 PM - 9 PM — HIPPI-6400

Tuesday, April 8 -
8 AM - 2 PM — HIPPI-TC General and -6400
2 PM - 5 PM — HIPPI-6400 Optical
5 PM - 9 PM — HIPPI-6400

The location is the Hyatt Regency Suites in Palm Springs, CA. Jeff Stai and Brocade are the host (See the meeting announcement on the web page at http://www.cic-5.lanl.gov/~det/ for further details.)

### 9.2 May 13-15, Mountain View, CA

This interim meeting will cover HIPPI-6400 and HIPPI-ST issues only.

Tuesday, May 13 – 1 PM - 9 PM
Wednesday, May 14 – 8 AM - 9 PM
Thursday, May 15 – 8 AM - 5 PM

The meetings will be held on the SGI campus, with Greg Chesson as the host. The meeting announcement will be put on the web page at http://www.cic-5.lanl.gov/~det/ as soon as it is available.

### 9.3 Tentative interim meeting, September 9-11, Albuquerque, NM

### 9.4 Tentative interim meeting, November 4-6, Dallas, TX

Los Alamos had earlier agreed to host a tentative September meeting, and Raytheon E-Systems a tentative November meeting. We had hoped that we would have the documents far enough along that the September and November meetings would not be necessary and could be canceled. At this meeting, we were still unsure if September and November meetings would be needed, but agreed to take Los Alamos and Raytheon E-Systems off the hook – the time required to set up a meeting in a hotel doesn't allow them enough time if we continue to delay the decision. Greg Chesson said that SGI would be able to host those meetings if they were deemed necessary, and they did not need as much lead time.

## 10. Future meeting dates and locations

The HIPPI, and X3T11, meeting schedule were reviewed.

The following 1997 X3T11 plenary week dates are firm. Recent changes to this list are underlined to make them easier to find.

**1997** –

| | | | |
|---|---|---|---|
| Apr 7-8 | Plenary | Palm Springs, CA | Brocade |
| May 13-15 | Interim | Mt. View, CA | SGI |
| Jun 9-10 | Plenary | Seattle, WA | Boeing |
| July 8-10 | Interim | Minneapolis, MN | Cray |
| Aug 4-5 | Plenary | Honolulu, HI | Hitachi |
| Sep 9-11 | Tentative | ?? | ?? |
| Oct 6-7 | Plenary | Tucson, AZ | FSI |
| Nov 4-6 | Tentative | ?? | ?? |
| Dec 8-9 | Plenary | Orlando, FL | DPT |

The 1998 schedule is less firm, but here is what is currently being considered by X3T11 for the plenary meetings. Question marks note the ones that are still in question. Hopefully HIPPI-6400 will be far enough along that we will not need interim working meetings.

**1998** –

| | | | |
|---|---|---|---|
| Feb 9-10 | Plenary | San Diego | Qlogic |
| Apr 20-21 | Plenary | Palm Springs, CA | Brocade |
| Jun 8-9 | Plenary | St. Petersburg Beach, FL | AMP |
| Aug 10-11 | Plenary | ?? | ?? |
| Oct 5-6 | Plenary | Ft. Lauderdale, FL | Adaptec |
| Dec 7-8 | Plenary | ?? | ?? |

## 11. Review action items

1. Greg Chesson to have Art Beckman look into getting a 24-bit organization identifier for HNF's use in ULAs.
2. Greg Chesson and Don Sanders to have SGI review the connector layout in HIPPI-6400-PH as it relates to the SuMAC chip.
3. Hansel Collins to obtain the driver output impedance values for HIPPI-6400-PH table 9.
4. Hansel Collins to check the shortest HIPPI-6400-PH cable length possible when using the on-board equalizer presently optimized for 50 meter cables.
5. John Ellis and Herb Van Deusen to investigate the possibility of putting the HIPPI-6400-PH passive equalizer networks within the cable assembly, e.g., connector backshell.
6. Don Tolmie to draft a Service Interface for HIPPI-6400-PH.
7. Don Tolmie to update HIPPI-6400-PH with the changes agreed to at the March meeting.
8. Don Tolmie to draft a Project Proposal and initial HIPPI-6400-OPT draft document.
9. Greg Chesson to propose text to replace the HIPPI-ST Notify bit. The intent is to change the polarity of the bit, i.e., to "Mute" or "Suppress" notification of DATA Operations.
10. Greg Chesson to propose text on the e-mail reflector describing HIPPI-ST Request_To_Receive set up using the Persistent bit.
11. Greg Chesson to draft initial text for HIPPI-ST Block acknowledge/retransmission functions, and consider how long an Originating Source should keep buffers for possible retransmission.
12. Greg Chesson and Jeffrey Chung to look at methods for rejecting a HIPPI-ST Request_To_Receive Operation.
13. Greg Chesson and Jeffrey Chung to consider developing "reason codes" to explain why a particular HIPPI-ST Operation was rejected.
14. Jeffrey Chung to verify correctness of HIPPI-ST summary tables 2 and 3 in relation to the text describing the individual Operations.
15. Jeffrey Chung to finish revising HIPPI-ST Annex C for the next revision of ST.
16. Greg Chesson and Jeffrey Chung to draft state tables for HIPPI-ST Operations.
17. Don Tolmie to update HIPPI-ST with the changes agreed to at the March meeting.
18. Craig Davidson to update his proposed HIPPI-ST annex for "Translating ST between HIPPI-6400 and HIPPI-FP" with the changes agreed to at the March meeting.
19. Michael McGowen to review 8.2.1 of HIPPI-6400-SC and propose a fix for when two directly connected end points try to use the same ULA.
20. Roger Ronald to update HIPPI-6400-SC with the changes agreed to at the March meeting.
21. Fred Templin to make a PDF copy of the fabric structure that was used in the HIPPI-6400-SC broadcast discussions at the March meeting. Don Tolmie to put the PDF copy on the web site as an aid for further discussions.
22. Fred Templin, Jeff Young, and Greg Chesson to begin an IP and ARP over HIPPI-6400 RFC.
23. Don Tolmie to allocate ULP-ids in HIPPI-FP for HIPPI-6400 style ARP, Reverse ARP, and IP.
24. Von Welch to contact HIPPI-6400 MIB developers and users for comments on current draft, and to prepare a presentation on the MIB for a future meeting.
25. Don Tolmie to provide an updated HIPPI meeting attendees list.

## 12. Adjournment

The group adjourned at 2 PM after a fruitful and intense meeting.

## 13. Attendance

| | | | |
|---|---|---|---|
| John Ellis | Berg | 717-938-7512 | Ellis825@aol.com |
| Austin Washington | Berg Electronics | 510-683-0700 | washington@harborcable.com |
| Michael McGowen | Essential Communications | 505-344-0080 | mikemc@esscom.com |
| Randy Hardy | Harris | 407-984-5654 | rhardy@harris.com |
| Chris Olson | Lockheed Martin | 602-925-7275 | olsonc@master.tds-az.lmco.com |
| Don Tolmie | Los Alamos National Lab | 505-667-5502 | det@lanl.gov |
| Fred Templin | NASA Ames | 415-604-6514 | templin@nas.nasa.gov |
| Von Welch | NCSA / U of Illinois | 217-244-3095 | vwelch@ncsa.uiuc.edu |
| Joe Parker | Optivision | 415-855-1775 | parker@optivision.com |
| Craig Davidson | Raytheon E-Systems | 214-205-4166 | davidson@esy.com |
| Bill McCoy | Raytheon E-Systems | 972-205-7234 | |
| Roger Ronald | Raytheon E-Systems | 972-205-8043 | rronald@esy.com |
| Greg Chesson | Silicon Graphics | 415-962-3496 | greg@sgi.com |
| Jeffrey Chung | Silicon Graphics | 415-933-3823 | jdchung@sgi.com |
| Hansel Collins | Silicon Graphics | 415-933-2921 | hac@engr.sgi.com |
| Bob Newhall | Silicon Graphics | 415-933-5424 | newhall@asd.sgi.com |
| Dave Parry | Silicon Graphics | 415-933-2192 | parry@asd.sgi.com |
| Don Sanders | Silicon Graphics | 415-933-6087 | dsanders@sgi.com |
| Chris Satterlee | Silicon Graphics | 415-933-2606 | csatt@sgi.com |
| Wally St.John | Silicon Graphics | 415-933-7613 | wbs@eng.sgi.com |
| Albert Kelley | Tensolite | 904-829-5600 | akelley@tensolite.com |
| Herb Van Deusen | W.L. Gore and Associates | 302-368-2575 | hvandeus@wlgore.com |